



HAL
open science

Parameter estimation for a class of stable driven stochastic differential equations

Solym Manou-Abi

► **To cite this version:**

Solym Manou-Abi. Parameter estimation for a class of stable driven stochastic differential equations. 2023. hal-04268224v3

HAL Id: hal-04268224

<https://hal.science/hal-04268224v3>

Preprint submitted on 28 Nov 2023 (v3), last revised 5 Dec 2023 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

PARAMETER ESTIMATION FOR A CLASS OF STABLE DRIVEN STOCHASTIC DIFFERENTIAL EQUATIONS

Solym M. Manou-Abi ^{1,2,3}

- ¹ Institut Montpellierain Alexander Grothendieck, UMR CNRS 5149, Montpellier
solym-mawaki.manou-abi@umontpellier.fr
- ² Centre Universitaire de Formation et de Recherche, Mayotte
Département Sciences et Technologies
solym.manou-abi@mayotte.fr
- ³ Laboratoire de Mathématiques et Applications, UMR CNRS 7348, Poitiers
solym.manou.abi@math.univ-poitiers.fr

ABSTRACT. In this paper, we study the estimation of diffusion parameters for one-dimensional, ergodic stochastic processes observed at some discrete times, that is a solution of a given class of stochastic differential equations driven by α -stable processes, $\alpha \in (1, 2)$. Firstly, we recall the non-parametric estimation framework of the drift coefficient namely the Nadaraya-Watson or Local polynomial estimators. Secondly, using the Euler-Maruyama scheme, we discuss an estimation method of the diffusion parameters (the scaling and the driving stable process parameters) based on a sample characteristic function combined with a Linear least squares estimation. We also consider the joint estimation of the drift and diffusion parameters for Stable driven Cox–Ingersoll–Ross (CIR) and Ornstein-Uhlenbeck (OU) processes. We discuss the validity and efficiency of the numerical implementation of the estimators using synthetic data. A real data in finance, such as exchange rates is used to fit the parameters of the Stable driven CIR or OU processes as a numerical estimation example. As a forthcoming work, we would like to establish proofs of the rate of convergence of our estimators in order to create a package on R software to handle this kind of estimation problem.

Key words and phrases: parameter estimation, Nadaraya-Watson estimation, stable process, characteristic function, regression method, CIR and OU processes.

1. INTRODUCTION

With the increasing of computational powerful statistical methods, there has been a great interest in parameter estimation for Stochastic Differential Equation (SDE). Such models are mathematical tools to describe the time evolution of many natural phenomena in many disciplines such as epidemiology, biology and finance. As examples, the Cox–Ingersoll–Ross (CIR) and the Ornstein-Uhlenbeck (OU) stochastic models [31, 9], has been used widely in finance. Some equivalent SDE models from continuous-time branching

processes are also shown to be useful models to study dynamic population [2, 39]. More precisely in epidemic dynamics, when an infectious population size is small and the population size is large, we can approximate the continuous time dynamic by a nonlinear SDE (stochastic logistic model that includes variability due to births and deaths [2]). For classical estimation methods, issues such as how to construct procedures which are both computationally efficient and show optimal statistical performance are now well understood. Let us consider the parameter estimation problem of the following SDE driven by a standard strictly stable process $(Z_t)_{t \geq 0}$ defined in a given filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$:

$$\begin{cases} dX_t = f(X_t)dt + \rho g(X_{t-})dZ_t, & t \in [0, T] \\ X_0 = x_0, \end{cases} \quad (1)$$

where x_0 is a starting point, $T > 0$ a given time horizon and the given function $g : \mathbb{R} \rightarrow \mathbb{R}$ is assumed to be known. We assume that, the function $f : \mathbb{R} \rightarrow \mathbb{R}$, constant $\rho > 0$ and the stable process $(Z_t)_{t \in [0, T]}$ (which is considered as a nuisance parameter) are unknown. The random variable Z_1 which is a standard strictly stable distribution with parameters $\alpha \in (1, 2)$ (the index of stability) and $\beta \in [-1, 1]$ (the skewness parameter) will be defined in the sequel. The existence and uniqueness of the solution to the SDE (??) under Lipschitz conditions are standard results in stochastic calculus [4]. For non Lipschitz coefficients, some results have been studied. Let us recall some few existence results. For $f = 0$ and g is Hölder continuous with exponent $\frac{1}{\alpha}$; if Z is a symmetric stable process then it is proved in [5] that there exists a strong solution for $\alpha \in (1, 2)$. If f and g have at most linear growth it is proved in [13] that there exists a weak solution for $\alpha \in (1, 2)$ for which pathwise uniqueness holds whenever the function g is Hölder continuous with exponent lying in $[1 - \frac{1}{\alpha}, \frac{1}{\alpha}]$. When the drift term f is Hölder continuous of order lying in $(2 - \alpha, 1)$ and g is Lipschitz continuous and bounded, then existence and uniqueness of a strong solution is derived in [35], [38]. For more recent existence conditions, we refer to [32] and references therein.

The main focus of this paper is the estimation of diffusion parameters for one-dimensional ergodic stochastic processes X observed at some discrete times, that is a solution of the SDE (1) driven by α -stable processes, $\alpha \in (1, 2)$. We assume that, the process X is observed at discrete time points $\{t_i = i\Delta_n, \quad i = 0, 1, 2, \dots, n\}$ with Δ_n a time frequency of the observation and n is the sample size. Firstly, we recall not only the non-parametric estimation framework of the drift coefficient but also the link between ergodicity and mixing conditions.

The parameter estimation theory for SDE driven by Brownian motions are well known in the literature. Some traditional methods are the maximum likelihood estimator (MLE) or the least squares estimator (LSE) techniques,

[29, 12, 24, 10], based on the Girsanov density. The consistency and the asymptotic distribution are well studied, see for instance [12, 24], [18], [40], [1]. For a more recent comprehensive discussion, we refer to [23] and the references therein. Substantial progress has been made in parameter estimation for SDE driven by Lévy processes with finite moments. The work in [34] dealt with the consistency and asymptotic normality when the driving process is a zero-mean adapted Lévy process with finite moments. The asymptotic normality of the LSE and MLE for the pure jump case is studied in [44, 45]. Note also that, some work has been carried out, when the driving process Z is a stable Lévy process which is really particular, due to its infinite variance property. However, the MLE is no longer valid in this setting because, the explicit density function is not always available and the Girsanov measure transformation is not well defined for α -stable processes, $\alpha \in (1, 2)$. Let us pointed out some well-known results. In [14, 15], authors use the trajectory fitting method combined with the weighted least squares technique of the drift coefficient for an α -stable driven OU process, $\alpha \in (1, 2)$, when the process is observed at discrete time instants, both for the ergodic and non ergodic case. They also discuss the consistency and asymptotic distribution of the estimator which has a higher order of convergence than in the classical Gaussian case. In [26, 48], authors study the drift parameter estimation of a stable driven Cox–Ingersoll–Ross (CIR) model (which is a special subcritical continuous state branching process with immigration). They derive the consistency and central limit theorems of the conditional least squares estimators and the weighted conditional least squares estimators of the drift parameters based on low frequency observations. In [11], a Lasso and Slope drift estimators for Lévy-driven Ornstein–Uhlenbeck processes were considered. A maximum likelihood type estimation of the drift and volatility constant coefficient parameters in stable driven CIR is studied in [50]. In a recent paper [6], author address the joint parameter estimation of the drift parameters, scaling parameter for the diffusion coefficient and the jump activity parameter such as the index of stability α and skewness parameter $\beta \in [-1, 1]$ from high-frequency observations of the stable CIR process on a fixed time period. Their methodology is based on the approximation of the conditional distribution. Now let us turn to non parametric estimation, an important field of statistics which consists of estimating an unknown function from a sample of data without first giving a formula. Many authors have investigated the non parametric estimation of the drift function f in the setting of diffusion's driven by Brownian motions under various conditions. The major statistical properties such as the consistency and rate of convergence of non parametric methods for the Nadaraya-Watson (N-W) estimators [36, 47], under independence with identically distribution and also weak dependence conditions such as mixing conditions. For a complete review of non parametric methods for diffusion processes; see the survey paper by [16]. In the stable and non parametric setting, some results are also studied in [49, 30] and [41, 51, 28]. But

the case of the diffusion part is much harder. We recall from [30] that, if the solution of SDE (1) is stationary (the stationary distribution having continuous density) and strongly mixing, the authors established a non parametric estimation of the drift function using the N-W estimator. The statistical properties such as the consistency and rate of convergence of the N-W estimators under dependence conditions such as mixing are developed in [30]. When the drift function is linear, the estimation both for discrete and continuous observations was studied in [15, 14, 27, 6, 50] including the stable driven CIR and OU processes. All these methods developed and applied within the framework of the stable driven CIR and OU processes are conditional dependent on the observation of the scaling ρ , the noise $Z = (Z_t)_{t \in [0, T]}$ parameters and time horizon T . Their performance depends strongly on the time frequency Δ_n . Although the parameter estimation for Stable driven SDE have been developed in recent years, few works (for instance [6] in the CIR case) deal with the joint parameter estimation both for the drift coefficient, scaling parameter, the diffusion coefficient parameters and the underlined standard strictly stable parameters such as the index of stability α and the skewness β .

This paper is organized as follows. After the introduction, we recall in Section 2, basic fact on stable processes, relevant assumptions for the N-W estimation framework as well as the connection between ergodicity and mixing conditions for Markov processes. In section 3, we state our main theorems, which consists on the explicit estimators formulas for the diffusion parameters namely the scaling parameter ρ and the stable noise parameters α and β . Using the Euler–Maruyama scheme and the N-W estimator of the drift coefficient, we discuss an estimation method based on a sample characteristic function combined with a Linear least squares estimation method. We also give estimators of the drift coefficients in the case of Stable driven CIR and OU processes as well as their consistency property. Section 4 is dedicated to the proofs of our main results. The last section investigate the numerical performance of the estimators on simulated data. A real data in finance, such as exchange rates of the Canadian dollar against the US dollar is used to fit the parameters of the best stable driven processes as a numerical estimation example. As a forthcoming work, we would like to establish the rate of convergence of some estimators in order to create a package on R software to handle this kind of estimation problem.

2. PRELIMINARIES

Throughout this paper we assume that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, P)$ is a filtered probability space satisfying the usual conditions, i.e.,

- (1) $(\Omega, \mathcal{F}_t, P)$ is complete for all $t \in \mathbb{R}_+$, \mathcal{F}_0 contains all the P-null sets in \mathcal{F} for all $t \in \mathbb{R}_+$.

- (2) $\mathcal{F}_t = \mathcal{F}_{t+}$ where $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$, for all $t \geq 0$, i.e. the filtration is right-continuous.

2.1. Stable processes and Nonparametric estimation of the drift function. In this section our objective is to introduce firstly, stable laws and processes. Secondly, we present a nonparametric estimation framework for the drift function of the SDE (1). Stable laws were introduced by Paul Lévy in 1925 and arise as the limit of normalized sums. The definitions and properties given below are taken from [42] and [37]. Note that there are multiple characterizations of stable distributions and each of them have some advantages depending on the purposes.

Definition 2.1. *A random variable X follows an α -stable distribution with $\alpha \in (0, 2]$ and we write $X \sim S_\alpha(\sigma, \beta, \omega; 1)$ if it is uniquely determined by its characteristic function :*

$$\Psi(t) = \mathbb{E}(\exp(itX)) = \begin{cases} \exp(-\gamma^\alpha |t|^\alpha [1 - i\beta(\tan(\frac{\pi\alpha}{2}))\text{sign}(t)] + i\omega t) & \text{if } \alpha \neq 1. \\ \exp(-\sigma |t| [1 + i\beta \frac{2}{\pi} \text{sign}(t) \log(|t|)] + i\omega t) & \text{if } \alpha = 1 \end{cases}$$

where $\alpha \in (0, 2]$ is the index of stability, $\beta \in [-1, 1]$ the skewness parameter; $\sigma > 0$ the scale parameter and $\omega \in \mathbb{R}$ the location or shift parameter and $\text{sign}(t)$ the sign function.

The above definition is the parametrization 1 in [37]. The following is the parametrization 0 and is usefull for computer processing.

Definition 2.2. *A random variable $X \sim S_\alpha(\sigma, \beta, \omega; 0)$ if,*

$$\Psi(t) = \begin{cases} \exp(-\sigma^\alpha |t|^\alpha [1 + i\beta(\tan(\frac{\pi\alpha}{2}))\text{sign}(t)(|\sigma t|^{1-\alpha} - 1)] + i\omega t) & \text{if } \alpha \neq 1. \\ \exp(-\sigma |t| [1 + i\beta \frac{2}{\pi} \text{sign}(t) \log(\gamma |t|)] + i\omega t) & \text{if } \alpha = 1. \end{cases}$$

For more detailed discussion on stable distributions, see [37]. When $\sigma = 1$ we say X is a standard α -stable random variable. For $\alpha \neq 1$ and $\omega = 0$ we say that X is a strictly α -stable random variable. Furthermore if $\beta = 0$ it is said to be symmetric. Note that, the parametrization 0 is important for numerical implementation because the characteristic function, density and cumulative distribution function are continuous with respect to the above four parameters α , β , γ and ω . Now let us introduce α -stable processes that can be seen as a particular class of Lévy processes and we refer to the well-known standard literature like [20].

Definition 2.3. *An \mathcal{F}_t -adapted stochastic process $Z = \{Z_t\}_{t \geq 0}$ is called an α -stable process if*

- (1) $Z_0 = 0$, a.s.;
- (2) Z has α -stable stationary increments distributions $Z_t - Z_s \sim Z_{t-s} \sim S_\alpha((t-s)^{1/\alpha} \sigma, \beta \omega)$, $t > s \geq 0$;
- (3) For any time points $0 \leq s_0 < \dots < s_m < \infty$, the random variables $Z_{s_0}, Z_{s_1} - Z_{s_0}, \dots, Z_{s_m} - Z_{s_{m-1}}$ are independent.

We also refer to [30] for the case of standard α -stable process. Now, we recall a non-parametric estimation framework of the drift coefficient of the following Stable driven SDE equation:

$$\begin{cases} dX_t = f(X_t)dt + \rho g(X_{t-})dZ_t, & t \in [0, T] \\ X(0) = x_0 \end{cases}$$

where f is an unknown function, $\rho > 0$ an unknown diffusion coefficient and the driving standard strictly stable process Z is unknown too. We assume in this paper that only the function g is known. Let X be a solution of the above SDE. Note that Lipschitz and Hölder (or bounded) conditions are typical conditions to show the existence of a solution. We refer to [32] and references therein. The Nadaraya-Watson (N-W) estimator is a classical method to estimate the drift function f . It is in this way that, we make the following assumptions following [30]. Similar results on the drift function estimation is also presented in [28] where authors study the local polynomial estimation under others regular conditions. We consider kernel function $K(\cdot)$ which is a symmetric and non negative probability density function satisfying $\sup(1 \vee |u|)K(u) < M_0 < +\infty$ and

$$\int_{-\infty}^{+\infty} u^2 K(u)du < +\infty, \quad \int_{-\infty}^{+\infty} K^2(u)du < +\infty.$$

- (**A₁**). The solution X_t is stationary and admits a unique invariant distribution π which is geometrically strong mixing.
 (**A₂**). The density function $f(x)$ of the stationary distribution π is continuous.
 (**A₃**). As $n \rightarrow \infty : h \rightarrow 0, \Delta_n \rightarrow 0,$

$$n\Delta_n \rightarrow \infty, \text{ and } \frac{n\Delta_n h}{(\log(n\Delta_n))^2} \rightarrow +\infty.$$

Assume that, the process X is observed at some discrete time points $\{t_k = k\Delta_n, k = 0, 1, 2, \dots, n-1\}$ with Δ_n a time frequency of the observation and n is the sample size. The Euler-Maruyama scheme of the above SDE is written as follows.

$$X_{t_{k+1}}^n = X_{t_k}^n + f(X_{t_k})\Delta_n + \rho g(X_{t_k})\Delta Z_k, \quad X_0^n = x \in \mathbb{R},$$

where $\Delta Z_k = Z_{t_{k+1}}^n - Z_{t_k}^n$. Set $Y_k = X_{t_{k+1}}^n - X_{t_k}^n$. The main idea of N-W estimator is to minimize the following object function:

$$\sum_{k=0}^{n-1} W_{n,k}(x)(Y_k - b\Delta_n)^2$$

over the set of parameters b with certain weights functions $W_{n,k}(x)$ given by:

$$W_{n,k}(x) = \frac{K_h\left(\frac{X_{t_k} - x}{h}\right)}{\sum_{k=0}^{n-1} K_h\left(\frac{X_{t_k} - x}{h}\right)}$$

for all $x \in \mathbb{R}$, where $K_h(x) = (K(x/h))/h$ with K the Kernel and h the bandwidth parameter. Thus the N-W estimator of the drift function f is given by the following expression :

$$\hat{f}_n(x) = \sum_{k=0}^{n-1} W_{n,k}(x) Y_k, \quad \forall x \in \mathbb{R}$$

so that

$$\hat{f}_n(x) = \sum_{i=1}^n \frac{K_h\left(\frac{X_{t_i}-x}{h}\right)(X_{t_i} - X_{t_{i-1}})}{n^{-1} \sum_{i=1}^n K_h\left(\frac{X_{t_i}-x}{h}\right)} \quad \text{for all } x \in \mathbb{R}. \quad (2)$$

We explore various Kernel functions in this paper but we choose only the one which perform the simulation. For instance we'll consider the Gaussian kernel function:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}.$$

2.2. Ergodic Markov processes and mixing conditions. We present in this section, some results concerning the link between ergodicity and mixing conditions for Markov processes. Let $(X_t, t \geq 0)$ be an ergodic Markov process with unique invariant measure π . It is well known that solution of some classical SDE driven by Lévy processes are Markov processes. Denote by $P_t(x, dz)$ be the transition probability of the process X_t starting from $X_0 = x$. Consider the Markov semigroup $(P_t)_{t \geq 0}$ associated to $(X_t, t \geq 0)$ and defined by $P_t f(x) = \mathbb{E}(f(X_t)/X_0 = x)$ for all measurable and positive functions or in $L^p(\pi)$. Recall that P is a bounded operator in all $L^p(\pi)$, $p \geq 1$ with operator norm equal to 1 (i.e. a contraction). Let us also recall the following adjoint operator P^* , defined by

$$\int f P_t g d\pi = \int g P_t^* f d\pi$$

for f and g square integrable w.r.t. π , which is again a contraction. The following definition introduces some way to control the ergodic decay to equilibrium.

Definition 2.4 (Ergodic rates of convergence [7]). *For any $r \geq p \geq 1$ and $t \geq 0$ we define the following ergodic rates*

$$\eta_{p,r}(t) = \sup_{\|f\|_{L^r(\pi)} \int f d\pi = 0} \|P_t f\|_{L^p(\pi)}.$$

The process X is said to be uniformly ergodic if $\lim_{t \rightarrow +\infty} \eta_{2,\infty}(t) = 0$ and then we call $\eta_{2,\infty}$ the uniform decay rate.

The following definitions can be found in [52].

Definition 2.5 (Exponentially or strongly ergodic process). *Assume that $X = (X_t, t \geq 0)$ is an ergodic Markov process with unique invariant measure π and $X_0 = x$.*

1. *The process X with is called exponentially ergodic if there exist a constant $k > 0$ and a positive measurable function $c(x)$ such that for all $t > 0$ we have*

$$\|P_t(x, \cdot) - \pi\|_{var} \leq c(x)e^{-kt},$$

2. *The process X with is called strongly ergodic if there exist two constants $k, C > 0$ such that for all $t > 0$ we have*

$$\|P_t(x, \cdot) - \pi\|_{var} \leq Ce^{-kt}.$$

where $\|\cdot\|_{var}$ denotes the total variation norm on the space of signed probability measures defined by

$$\begin{aligned} \|P_t(x, \cdot) - \pi\|_{var} &= \sup_{A \in \mathcal{F}} |P_t(x, A) - \mu(A)| \\ &= \sup_{\|f\|_{\infty} \leq 1 \text{ and } Law(Y) = \pi} |\mathbb{E}f(X_t) - \mathbb{E}f(Y)|. \end{aligned}$$

From this definition, one say that the process X is said to be strongly or exponentially ergodic if $\lim_{t \rightarrow +\infty} \eta_{1,\infty}(t) = 0$ and the we call $\eta_{1,\infty}$ the strong or exponential ergodic decay rate. Now, let us define some usual mixing coefficients in order to compare mixing properties and ergodic decay properties.

Definition 2.6 (Mixing conditions, see [33]). *Let \mathcal{F}_s (resp. \mathcal{G}_s) be respectively the backward (or the past) and the forward (or the future) σ -fields generated by X_u for $0 \leq u \leq s$ (resp. $u \leq s$).*

1. *The strong mixing coefficient $\alpha_{mix}(t)$ is defined as :*

$$\begin{aligned} \alpha_{mix}(t) &= \sup_s \sup_{A \in \mathcal{F}_s, B \in \mathcal{G}_{s+t}} |\mathbb{P}(A \cup B) - \mathbb{P}(A)\mathbb{P}(B)| \\ &= \frac{1}{4} \sup_s \left\{ \sup_{F, G} Cov(F, G), F \in \mathcal{F}_s \text{ (resp. } G \in \mathcal{G}_{s+t}) \text{ measurable and bounded by } 1. \right\} \end{aligned}$$

If $\lim_{t \rightarrow \infty} \alpha_{mix}(t) = 0$, the process is strongly mixing.

2. *The β -mixing coefficient $\varphi(t)$ is defined as:*

$$\beta_{mix}(t) = \sup_s \left\{ \sup_{A, B} (\mathbb{P}(B|A) - \mathbb{P}(B)), A \in \mathcal{F}_s, B \in \mathcal{G}_{s+t} \right\}.$$

If $\lim_{t \rightarrow \infty} \beta_{mix}(t) = 0$, the process is β -mixing or uniformly mixing.

3. The ρ -mixing coefficient $\rho_{mix}(t)$ is defined as the maximal correlation coefficient, i.e.

$$\rho_{mix}(t) = \sup_s \left\{ \sup_{F,G} \text{Corr}(F,G), F \in L^2(\mathcal{F}_s), G \in L^2(\mathcal{G}_{t+s}) \right\}.$$

If $\lim_{t \rightarrow \infty} \rho_{mix}(t) = 0$ the process is ρ -mixing.

Note that if X is a stationary process (i.e. such that, for all $n \geq 0$, the law of X_{t+s} is the same as the one of X_s), the supremum on s is irrelevant. The following Lemma is contained in our past paper [7] and allow to make a connection between ergodicity and mixing conditions.

Lemma 2.1. *For all $t \geq 0$, we have*

- (1) $\eta_{\infty,2}^2(t) \vee (\eta^*)_{\infty,2}^2(t) \leq \alpha_{mix}(t) \leq \eta_{\infty,2}([t/2]) \eta_{\infty,2}^*(t/2)$.
- (2) Either $\eta_{2,2}(t) = 1$ for all t or $\eta_{2,2}(t) \leq c e^{-\lambda t}$ for some $\lambda > 0$. In the second case

$$\eta_{2,2}^2(t) = (\eta^*)_{2,2}^2(t) \leq \rho_{mix}(t) \leq c \eta_{2,2}(t).$$

- (3) $\beta_{mix}(t) \leq \eta_{1,\infty}^2(t/2)$.

From the above Lemma one easily derive the following result.

Theorem 2.2. *Assume that $X = (X_t, t \geq 0)$ is an ergodic Markov process with unique invariant measure π .*

- a) *If X is strongly or exponentially ergodic then it is β -mixing.*
- b) *If X is uniformly ergodic, then it is strongly mixing.*
- c) *Any kind of exponential ergodic decay rate in L^2 imply the ρ -mixing.*

3. MAIN RESULTS

Using the N-W estimator and the Euler–Maruyama scheme, the main objective in this section is firstly the joint estimation of the scaling parameter ρ and the standard strictly stable process parameters α and β of the diffusion component in the SDE in (2.1).

Theorem 3.1. *Let $X = (X_t)_{t \in [0,1]}$ be a real valued stationnary and strongly mixing process with unique invariant distribution which is solution of the SDE (2.1). We assume that X is observed on a sample of discrete time points t_i with step Δ_n of size n such $g(X_{t_i}) \neq 0$. Assume that assumptions **(A₂)** and **(A₃)** holds. Let $(\hat{f}_n(X_{t_i}))_{i=1}^n$ be the sequence of the N-W estimate of the drift coefficient of the SDE (2.1). We define the following sample characteristic function*

$$\hat{\varphi}_n(u_k) = \frac{1}{n} \sum_{i=0}^{n-1} \exp \left(j u_k \left(\frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i}) \Delta_n}{g(X_{t_i})} \right) \right) \quad (3)$$

for all $k \in [[1, m]]$, $m > 0$, and $j^2 = -1$. Set

$$\left\{ \begin{array}{l} \hat{\alpha}_m = \frac{\sum_{k=1}^m W_k V_k - \frac{1}{m} \sum_{k=1}^m V_k \sum_{k=1}^m W_k}{\sum_{k=1}^m W_k^2 - \frac{1}{m} \left(\sum_{k=1}^m W_k \right)^2}, \\ \hat{\lambda}_m = \frac{1}{m} \sum_{k=1}^m V_k - \frac{\hat{\alpha}_m}{m} \sum_{k=1}^m W_k \\ \hat{\rho}_m = \Delta_n^{-\frac{1}{\hat{\alpha}_m}} \exp \frac{\hat{\lambda}_m}{\hat{\alpha}_m} \\ \hat{\beta}_m = \frac{\sum_{k=1}^m u_k S_k}{\sum_{k=1}^m u_k B_k}, \end{array} \right. \quad (4)$$

where

$$S_k = \arg(\hat{\varphi}_n(u_k)), \quad B_k = \tan\left(\frac{\pi \hat{\alpha}}{2}\right) \text{sign}(u_i) (u_k - u_k^{\hat{\alpha}_m}) \Delta_n^{1/\hat{\alpha}_m},$$

and

$$V_k = \log(-\log|\hat{\varphi}_n(u_k)|) \quad \text{and} \quad W_k = \log(|u_k|).$$

For n large enough, $\hat{\alpha}_m$, $\hat{\rho}_m$ and $\hat{\beta}_m$ are least squares estimators of α , ρ and β .

Let's discuss about some formulations in order to check the consistency through a linear statistical regression model.

Remark 3.1. *Depending on the normality for the observed distributions of the sample characteristic function $\hat{\varphi}_n$ estimation from the N - W estimator; the above estimators can be consistent and unbiased if the errors have finite variance and are uncorrelated with the regressors $\log(|u_k|)$. The condition that the errors are uncorrelated with the regressors is here satisfied in an experimentation since they are deterministic. If the sample prediction errors follow a Gaussian distribution then the above estimators are consistent and unbiased. If this is not the case, note that data normalization transformations are available in order to have Gaussian observed distributions of the sample characteristic function. For instance, the Min-Max Scaling can be applied since the data may vary in different scales, reducing the effect of outliers.*

A simulation study is conducted to evaluate the performance of the proposed estimation. The choice of optimal value of m was proposed in [21], which suggest selecting points u_i in the interval $[0.1, 1]$. We give the proof of Theorem 3.1 in section 4.

Theorem 3.1 also applies to the following α -stable-driven Ornstein Uhlenbeck (OU) and Cox–Ingersoll–Ross (CIR) processes. Such models are popular in stochastic modelling for description of interest rates in finance and population dynamics. Assume that Z is a standard strictly stable process with parameters $\alpha \in (1, 2)$ and $\beta \in [-1, 1]$. A generalized Ornstein-Uhlenbeck (OU) process driven by an α -stable process $(Z_t)_{t \geq 0}$ is defined to be the solution to the following linear stochastic differential equation

Definition 3.1. (*Stable OU process*)

$$dX_t = \theta(\mu - X_t)dt + \rho dZ_t, \quad X_0 = x \in \mathbb{R}. \tag{5}$$

and has the integral representation for any $t \geq r \geq 0$:

$$X_t = e^{-\theta(t-r)}X_r + \theta\mu \int_r^t e^{-\theta(t-s)}ds + \rho \int_r^t e^{-\theta(t-s)}dZ_s, \quad t \geq 0.$$

where θ and μ are constants.

The stable driven Cox–Ingersoll–Ross (SCIR [50]) is defined by:

$$dX_t = \theta(\mu - X_t)dt + \rho|X_t|^{1/q}dZ_t, \quad X_0 \geq 0, \tag{6}$$

where $q > 0$, $\theta > 0$, $\rho > 0$, and $\mu \in \mathbb{R}$ are constants and Z is an α -stable process. In the case of a pure-jump α -stable Lévy process Z with positive jumps the following extension is introduced in the literature (see [6]) given as follows for $q = \alpha$ when $\alpha \in (1, 2)$.

Definition 3.2. (*Stable CIR process*) An α -stable process is defined by:

$$dX_t = \theta(\mu - X_t)dt + \rho|X_t|^{1/\alpha}dZ_t, \quad X_0 \geq 0,$$

where $\theta > 0$, $\rho > 0$, and $\mu \in \mathbb{R}$ are constants.

This stable driven CIR model is a particular form of the so-called continuous-state branching processes with immigration (CBI-processes, [25], [26] and [39]), which arise as scaling limits of Galton–Watson branching processes with immigration (GWI-processes); see, e.g., [19].

The following result discuss the parameter estimation of the drift coefficients θ and μ of the above stable driven CIR and OU processes.

Theorem 3.2. *Let $X = (X_t)_{t \in [0,1]}$ be a real valued stationary and strongly mixing process satisfying (5) or (6). We assume that X is observed on a sample of discrete time points t_i with time frequency Δ_n of size n . Let $(\hat{f}_n(X_{t_i}))_{i=1}^n$ be the sequence of the N-W estimate of the linear drift function*

in (5) or (6) and set

$$\left\{ \begin{array}{l} \hat{a}_n = \frac{\sum_{i=0}^{n-1} X_{t_i} \hat{f}_n(X_{t_{i+1}}) - \frac{1}{n} \sum_{i=0}^{n-1} X_{t_i} \sum_{i=1}^n \hat{f}_n(X_{t_{i+1}})}{\frac{1}{n} \left(\sum_{i=0}^{n-1} X_{t_i} \right)^2 - \sum_{i=0}^{n-1} X_{t_i}^2} \\ \hat{\mu}_n = \hat{a}_n^{-1} \frac{1}{n} \sum_{i=0}^{n-1} \hat{f}_n(X_{t_{i+1}}) + \frac{1}{n} \sum_{i=0}^{n-1} X_{t_i} \\ \hat{\theta}_n = -W(\hat{a}_n), \end{array} \right. \quad (7)$$

where W is the Lambert function.

- We assume moreover that the driving process Z in the stable driven CIR process (6) is an α -stable symmetric process with $\alpha \in (\sqrt{2}, 2)$.

For n large enough, $\hat{\theta}_n$ and $\hat{\mu}_n$ are consistent estimators of θ and μ .

4. PROOFS OF THE MAIN THEOREMS

In this section, we derive the proofs of our main results.

4.1. Proof of Theorem 3.1.

Proof. We recall first the Euler Euler–Maruyama approximation of the SDE (1):

$$X_{t_i}^n = X_{t_{i-1}}^n + f(X_{t_i}^n) \Delta_n + g(X_{t_i}^n) Y_i^n$$

where $Y_i^n = \rho \Delta Z_i^n = \rho (Z_{t_{i+1}}^n - Z_{t_i}^n)$, $i = 0, \dots, n-1$, are identically sequence distributed as $Z_{\Delta_n} \sim S_\alpha(\Delta_n^{1/\alpha} \rho, \beta, 0)$. We have

$$Y_i^n = \frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i}) \Delta_n}{g(X_{t_i})}, \quad \forall i = 0, \dots, n-1.$$

We can define an approximate sample characteristic function by

$$\hat{\varphi}_n(u) = \frac{1}{n} \sum_{i=0}^{n-1} \exp\left\{ju \left(\frac{\Delta X_{t_i} - \hat{f}_n(X_{t_i}) \Delta_n}{g(X_{t_i})} \right)\right\}, \quad \forall u \in \mathbb{R}, \quad (8)$$

which shall be asymptotically equal to the characteristic function $\varphi_{Z_{\Delta_n}}(u)$ of Z_{Δ_n} . Now, using Definition 2.1 or 2.2, we have the following formula:

$$\log(-\log|\varphi_{Z_{\Delta_n}}(u)|) = \log(\rho^\alpha \Delta_n) + \alpha \log(|u|).$$

From this, we can set up an ordinary least squares regression method in order to obtain the estimators of α and ρ . To this end, it is desired to find the vector parameter (α, λ) such that the underlined linear function fits best the given data in the least squares sense, that is, the following sum of squares :

$$G(\lambda, \alpha) = \sum_{k=1}^m (V_k - \lambda - \alpha W_k)^2$$

$$(\hat{\lambda}, \hat{\alpha}) = \arg \min_{(\lambda, \alpha)} G(\lambda, \alpha),$$

is minimized, where

$$V_k = \log(-\log|\hat{\varphi}_n(u_k)|), \quad W_k = \log(|u_k|), \quad \lambda = \log(\rho^\alpha \Delta_n).$$

The minimum value of $G(\lambda, \alpha)$ occurs when the gradient is zero. Since the model contains two parameters, there are two gradient equations:

$$\begin{cases} \frac{\partial G(\lambda, \alpha)}{\partial \alpha} = -2 \sum_{k=1}^m W_k (V_k - \lambda - \alpha W_k) \\ \frac{\partial G(\lambda, \alpha)}{\partial \lambda} = -2 \sum_{k=1}^m (V_k - \lambda - \alpha W_k), \end{cases}$$

and these gradient equations have a closed solution given by:

$$\begin{cases} \frac{\partial G(\lambda, \alpha)}{\partial \alpha} = 0 \\ \frac{\partial G(\lambda, \alpha)}{\partial \lambda} = 0 \end{cases} \implies \begin{cases} \hat{\alpha}_m = \frac{\sum_{k=1}^m W_k V_k - \frac{1}{m} \sum_{k=1}^m V_k \sum_{k=1}^m W_k}{\sum_{k=1}^m W_k^2 - \frac{1}{m} \left(\sum_{k=1}^m W_k \right)^2}, \\ \hat{\lambda}_m = \frac{1}{m} \sum_{k=1}^m V_k - \frac{\hat{\alpha}_m}{m} \sum_{k=1}^m W_k. \end{cases} \quad (9)$$

To estimate β , we use Definition 2.2. to obtain the following formula:

$$S_k = \beta B_k + \omega u_k \quad \text{here we have, } \omega = 0$$

since we consider in this work strictly α -stable process and

$$S_k = \arg(\hat{\varphi}_n(u_k)), \quad B_k = \tan\left(\frac{\pi \hat{\alpha}}{2}\right) \text{sign}(u_i) (u_k - u_k^{\hat{\alpha}}) \Delta_n^{1/\hat{\alpha}}.$$

We use again the above least squares method to minimize the following sum of squares:

$$T(\beta, \omega) = \sum_{k=1}^m (S_k - \beta B_k - \omega u_k)^2.$$

From the gradient equations:

$$\begin{cases} \frac{\partial T(\beta, \omega)}{\partial \beta} = -2 \sum_{k=1}^m B_k (S_k - \omega u_k - \beta B_k) \\ \frac{\partial T(\beta, \omega)}{\partial \omega} = -2 \sum_{k=1}^m u_k (S_k - \omega u_k - \beta B_k), \end{cases}$$

we have the following closed solution when $\omega = 0$:

$$\hat{\beta}_m = \frac{\sum_{k=1}^m u_k S_k}{\sum_{k=1}^m u_k B_k}.$$

□

4.2. Proof of Theorem 3.4.

Proof. Note that, by applying Itô's formula to (5) or (6) we have

$$X_{t_i} = \mu(1 - e^{-\theta}) + e^{-\theta} X_{t_{i-1}} + \epsilon_i$$

where

$$\epsilon_i = \begin{cases} \rho \int_{t_{i-1}}^{t_i} e^{-\theta(i-s)} |X_{s-}|^{1/\alpha} dZ_s & \text{for the stable CIR process} \\ \rho \int_{t_{i-1}}^{t_i} e^{-\theta(i-s)} dZ_s & \text{for the stable OU process.} \end{cases}$$

The sequence (ϵ_i) is a centered random variable (martingale differences) in the case of the stable OU process (5) if Z is a strictly α -stable process with $\alpha \in (1, 2)$ for with $\alpha \in (1, 2)$. If Z is a symmetric α -stable process and in the case of the stable CIR process (6), then according to the existence result in [5] and integrability property in [13] and [17] the sequence (ϵ_i) is a centered random variable with $\alpha \in (\sqrt{2}, 2)$ since $x \rightarrow |x|^{1/\alpha}$ is Hölder continuous with exponent α^{-1} . Therefore we have

$$\hat{f}(X_{t_{i+1}}) = a\mu - aX_{t_i} + \eta_i, \quad a = \theta e^{-\theta}, \quad i \geq 0,$$

where the sequence (η_i) is again a centered random sequence. We can set up a Linear Least Square Method which consists in minimizing the following objective function:

$$\mathbb{G}(a, \mu) = \sum_{i=0}^{n-1} (\hat{f}(X_{t_{i+1}}) - a\mu + aX_{t_i})^2.$$

Using, the gradient equations:

$$\begin{cases} \frac{\partial \mathbb{G}(a, \mu)}{\partial a} = 2(-\mu + X_{t_i}) \sum_{i=0}^{n-1} (\hat{f}(X_{t_{i+1}}) - a\mu + aX_{t_i}) \\ \frac{\partial \mathbb{G}(a, \mu)}{\partial \mu} = -2a \sum_{i=0}^{n-1} (\hat{f}(X_{t_{i+1}}) - a\mu + aX_{t_i}), \end{cases}$$

it is straightforward to obtain the following least square estimators :

$$\left\{ \begin{array}{l} \hat{a}_n = \frac{\sum_{i=0}^{n-1} X_{t_i} \hat{f}_n(X_{t_{i+1}}) - \frac{1}{n} \sum_{i=0}^{n-1} X_{t_i} \sum_{i=1}^n \hat{f}_n(X_{t_{i+1}})}{\frac{1}{n} \left(\sum_{i=0}^{n-1} X_{t_i} \right)^2 - \sum_{i=0}^{n-1} X_{t_i}^2} \\ \hat{\mu}_n = \hat{a}_n^{-1} \frac{1}{n} \sum_{i=0}^{n-1} \hat{f}_n(X_{t_{i+1}}) + \frac{1}{n} \sum_{i=0}^{n-1} X_{t_i} \\ \hat{\theta}_n = -W(\hat{a}_n), \end{array} \right. \quad (10)$$

where W is the Lambert function [8]. Now, let us prove the consistency of these estimators. Set $U_i = \hat{f}_n(X_{t_{i+1}})$ and note that since we assume that the process X is ergodic and $\alpha \in (1, 2)$ then $\bar{X} = \frac{1}{n} \sum_{i=0}^{n-1} X_{t_i}$ and $\bar{U} = \frac{1}{n} \sum_{i=0}^{n-1} U_i$ converge. We have

$$\hat{a}_n = \frac{\sum_{i=1}^n X_{t_i} U_i - \frac{1}{n} \sum_{i=1}^n U_i \sum_{i=1}^n X_{t_i}}{\frac{1}{n} \sum_{i=1}^n X_{t_i} \sum_{i=1}^n X_{t_i} - \sum_{i=1}^n X_{t_i}^2} = \frac{\sum_{i=1}^n X_{t_i} (U_i - \bar{U})}{\sum_{i=1}^n X_{t_i} (\bar{X} - X_{t_i})} \quad (11)$$

$$\text{where } \bar{U} = \frac{1}{n} \sum_{i=1}^n U_i \text{ and } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_{t_i} \quad (12)$$

so that

$$= \frac{\sum_{i=1}^n (X_{t_i} - \bar{X}) (\bar{U} - U_i)}{\sum_{i=1}^n (X_{t_i} - \bar{X})^2} \text{ since } \sum_{i=1}^n \bar{X} (U_i - \bar{U}) = 0 \text{ and } \sum_{i=1}^n \bar{X} (X_{t_i} - \bar{X}) = 0. \quad (13)$$

From the relation

$$U_i - U = a(\bar{X} - X_{t_i}) + \varepsilon_i,$$

where ε_i is a centered random sequence (martingale differences), we have

$$\hat{a}_n = \frac{\sum_{i=1}^n (X_{t_i} - \bar{X}) (a(X_{t_i} - \bar{X}) - \varepsilon_i)}{\sum_{i=1}^n (\bar{X} - X_{t_i})^2} = a + \frac{\sum_{i=1}^n (X_{t_i} - \bar{X}) \varepsilon_i}{\sum_{i=1}^n (\bar{X} - X_{t_i})^2} \xrightarrow{n \rightarrow +\infty} 0, \quad (14)$$

since the process $X = \{X_t, t \geq 0\}$ is ergodic, stationary and have infinite square variation. We deduce the consistency of $\hat{\theta}_n$ and $\hat{\mu}_n$ by the use of continuous mapping theorem. \square

5. NUMERICAL APPLICATION WITH STRONG ERGODIC STOCHASTIC MODELS

The numerical implementation was done with R software. For practical setting, we choose the optimal bandwidth h (appearing in the Gaussian Kernel for the N-W estimation) by considering the method proposed by [43]. The R function `h.amize` allow to do it. Note also that, in many cases people usually choose $h = h_n = n^{-1/5}$.

We now investigate the numerical performance of our estimation procedure on simulated data generated from ergodic stationary processes. The exponential ergodicity of a Lévy driven OU process is established in [46]. The result entails that if $\theta > 0$ then the α -Stable OU process (5) has a unique invariant measure π and it is strongly ergodic (mixing). More generally, let us consider the case where the drift f and diffusion coefficient g verify sufficient conditions for the solution of (1) to exist [5, 13]. According to [22], if $f(\cdot)$ is locally Lipschitz and $\limsup_{|x| \rightarrow +\infty} \frac{f(x)}{x} < 0$ then, for any $\alpha \in (1, 2)$, the solution of (1) is exponentially ergodic and its invariant distribution exists and is unique. This therefore applies to the case of the stable OU model (5) under the condition $\theta > 0$. In the case of the Stable driven CIR-model 6, some results exists in [26] whenever the stable process has only positive jumps, where for $\theta > 0$ and $\mu \geq 0$ it can be seen as a subcritical CBI process with immigration rate μ . Thus, using the result of [26], we conclude that the Stable driven CIR process is exponentially ergodic and hence strongly mixing.

5.1. Simulated data case's. We simulate and approximate X by using the Euler scheme on the interval $[0, T]$ with sample size $n = 500, 1000, 2000$ and within a period $T = 1$. In order to ensure that the Stable driven OU and CIR processes satisfies Assumption **(A₁)**, we shall employ the exponential ergodicity results we discuss above when choosing the model parameters. In the sequel, we consider the following common parameters:

$$x_0 = 1, \quad \theta = 0.3, \quad \mu = 1.2 \quad \alpha = 1.6 \quad \beta = 0 \quad \text{and} \quad \rho = 1.$$

We proceed by generating data from models (5) and (6) according to the previous setting with a strictly standard α -stable process. We plot the kernel density estimate of a realization of X with the histogram in order to check numerically the regularity condition on the density function of the stationary distribution π , assumed to be continuous in Assumption **(A₂)**. Therefore choosing a suitable time frequency (for instance $\Delta_n = \frac{1}{\sqrt{n}}$), the models satisfies all the required conditions (including Assumption **(A₃)**). We explore the performance of the Nadaraya-Watson estimator of the drift function and the

estimate parameters of the diffusion part, namely α , β and ρ in the context of the stable driven CIR and OU models on simulated data. Figures 1 and 2 display the graphical performance of the N-W estimator with respect to the true drift function considered ($\theta = 0.3$, $\mu = 1.2$). The pictures show the comparison according to three sample sizes $n = 500$, $n = 1000$ and $n = 2000$ with time interval $T = 1$ or $T = 10$. We can see that varying T for a fixed n slightly seems not improve the estimates of the drift function in the SCIR model. This confirms that the drift function can not be identified in a fixed time interval in the framework of N-W estimator, no matter how frequently the observations are sampled. One notices that for a fixed length of observation interval T , as the sample size gets larger, the N-W estimator does not behave better, which is consistent with the asymptotic theory of the N-W estimator for stochastic processes driven by Lévy motions. The regularity condition on the density function of the stationary distribution is graphically checked throughout Figures 3 and 4.

To evaluate the performance, we make not only illustration by picture but also, via the square-Root of Average Square Errors defined by

$$\text{RMSE} = \frac{1}{n} \sum_{i=1}^n \left(\hat{f}_n(x_i) - f(x_i) \right)^2$$

where $x = (x_i)$ are chosen uniformly to cover the range of sample path of X . We summarized the performance in Tables 1 and 4, which reports the results on the diffusion parameters (the standard and strictly stable process parameters α and β) and the scaling parameter ρ as well as the RMSE on n replicates with three sample sizes $n = 500$, $n = 1000$ and $n = 2000$, respectively for time interval $T = 1$ and $T = 10$. We can see that varying T for a fixed n slightly changes the estimates of the stable process parameters α, β , except for the scaling parameter ρ .

5.2. Real data case's: financial exchanges rates. We now turn to the estimation using real data. We analyse real financial data such as indexes or exchanges of the Canadian dollar against the US dollar in a given fixed period varying from May 2018 to June 2022 are analysed with a sample size $n = 1000$. Figures 5(a), 5(b) and 5(c) show the time series variation of the exchanges rates. We also look at the speed of decreasing of the Covariance using the auto-correlation function (ACF) diagram [3]. Indeed for a stationary time series, the ACF will fall to zero quickly whereas the ACF of non-stationary time series decreases slowly. The Dickey Fuller test returns a p-criticism of 0.01, which confirms the stationarity of the times series. The decreasing of the covariance using the ACF diagram test shows that the data sequence is weakly dependent. The data can be considered as mixing in the sense of the ρ -mixing (which implies strong mixing). The density estimation of the stationary distribution is given also in Figures 5 (a'), 5(b') and 5(c') so that the regularity condition is satisfied. Figure 6 display the graphical

N-W estimation of the observed drift function. Tables 3 and 5 summarise the parameter estimation of the diffusion part. As concerned the drift part, Tables 2 summarise the estimated drift coefficients with the real data. We compare the prediction result with the stable OU model in Figure 7 since it seems to match better than the Stable CIR process.

FUNDING

No funding was obtained for this study.

COMPETING INTEREST

The authors declare that they have no competing interests.

DECLARATION STATEMENT

The author certify that the submission is an original work and is not under review at any other publication.

DATA AVAILABILITY STATEMENTS

The real dataset used in this work is available online and free of charge. The R codes for the simulation study are available from the author on reasonable request.

REFERENCES

- [1] Yacine Aït-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, 70(1):223–262, 2002.
- [2] Linda JS Allen. Stochastic population and epidemic models. *Mathematical biosciences lecture series, stochastics in biological systems*, page 128, 2015.
- [3] Alexander Alvarez and Pablo Olivares. Méthodes d'estimation pour des lois stables avec des applications en finance. *Journal de la société française de statistique*, 146(4):23–54, 2005.
- [4] David Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.
- [5] Richard F Bass, Krzysztof Burdzy, and Zhen-Qing Chen. Stochastic differential equations driven by stable processes for which pathwise uniqueness fails. *Stochastic processes and their applications*, 111(1):1–15, 2004.
- [6] Elise Bayraktar and Emmanuelle Clément. Estimation of a pure-jump stable cox-ingersoll-ross process. *arXiv preprint arXiv:2304.02386*, 2023.
- [7] Patrick Cattiaux and S. Mawaki Manou-Abi. Limit theorems for some functionals with heavy tails of a discrete time markov chain. *ESAIM: Probability and Statistics*, 18:468–482, 2014.
- [8] Robert M Corless, Gaston H Gonnet, David EG Hare, David J Jeffrey, and Donald E Knuth. On the lambert w function. *Advances in Computational mathematics*, 5:329–359, 1996.
- [9] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. In *Theory of valuation*, pages 129–164. World Scientific, 2005.

- [10] Peter Craigmile, Radu Herbei, Ge Liu, and Grant Schneider. Statistical inference for stochastic differential equations. *Wiley Interdisciplinary Reviews: Computational Statistics*, 15(2):e1585, 2023.
- [11] Niklas Dexheimer and Claudia Strauch. On lasso and slope drift estimators for lévy-driven ornstein–uhlenbeck processes. *arXiv preprint arXiv:2205.07813*, 2022.
- [12] A Ja Dorogovcev. The consistency of an estimate of a parameter of a stochastic differential equation. *Theory Probab. Math. Stat.*, 10:73–82, 1976.
- [13] Nicolas Fournier. On pathwise uniqueness for stochastic differential equations driven by stable lévy processes. In *Annales de l'IHP Probabilités et statistiques*, volume 49, pages 138–159, 2013.
- [14] Yaozhong Hu and Hongwei Long. Parameter estimation for ornstein-uhlenbeck processes driven by stable lévy motions. *Communications on Stochastic Analysis*, 1(2):1, 2007.
- [15] Yaozhong Hu and Hongwei Long. Least squares estimator for ornstein–uhlenbeck processes driven by α -stable motions. *Stochastic Processes and their applications*, 119(8):2465–2480, 2009.
- [16] Stefano M Iacus et al. *Simulation and inference for stochastic differential equations: with R examples*, volume 486. Springer, 2008.
- [17] Aldéric Joulin and Solym Mawaki Manou-Abi. A note on convex ordering for stable stochastic integrals. *Stochastics An International Journal of Probability and Stochastic Processes*, 87(4):592–603, 2015.
- [18] RA Kasonga. The consistency of a non-linear least squares estimator from diffusion processes. *Stochastic processes and their applications*, 30(2):263–275, 1988.
- [19] Kiyoshi Kawazu and Shinzo Watanabe. Branching processes with immigration and related limit theorems. *Theory of Probability & Its Applications*, 16(1):36–54, 1971.
- [20] Sato Ken-Iti. *Lévy processes and infinitely divisible distributions*, volume 68. Cambridge university press, 1999.
- [21] Stephen M Kogon and Douglas B Williams. Characteristic function based estimation of stable distribution parameters. *A practical guide to heavy tails: statistical techniques and applications*, pages 311–338, 1998.
- [22] Alexey M Kulik. Exponential ergodicity of the solutions to sde's with a jump noise. *Stochastic Processes and their Applications*, 119(2):602–632, 2009.
- [23] Yu A Kutoyants. *Statistical inference for ergodic diffusion processes*. Springer Science & Business Media, 2004.
- [24] A Le Breton. On continuous and discrete sampling for parameter estimation in diffusion type processes. In *Stochastic Systems: Modeling, Identification and Optimization, I*, pages 124–144. Springer, 2009.
- [25] Pei-Sen Li, Zenghu Li, Jian Wang, and Xiaowen Zhou. Exponential ergodicity of branching processes with immigration and competition. *arXiv preprint arXiv:2205.15499*, 2022.
- [26] Zenghu Li and Chunhua Ma. Asymptotic properties of estimators in a stable cox–ingersoll–ross model. *Stochastic Processes and their Applications*, 125(8):3196–3233, 2015.
- [27] Zenghu Li and Leonid Mytnik. Strong solutions for stochastic differential equations with jumps. In *Annales de l'IHP Probabilités et statistiques*, volume 47, pages 1055–1067, 2011.
- [28] ZhengYan Lin, YuPing Song, and JiangSheng Yi. Local linear estimator for stochastic differential equations driven by stable lévy motions. *Science China Mathematics*, 57:609–626, 2014.
- [29] Robert Sevilevic Lipcer, Robert S Liptser, Albert Nikolaevic Shiraev, Albert N Shiraev, et al. *Statistics of Random Processes II: II. Applications*, volume 2. Springer Science and Business Media, 2001.

- [30] Hongwei Long and Lianfen Qian. Nadaraya-Watson estimator for stochastic processes driven by stable Lévy motions. *Electronic Journal of Statistics*, 7(none):1387 – 1418, 2013.
- [31] Ross A Maller, Gernot Müller, and Alex Szimayer. Ornstein–uhlenbeck processes and extensions. *Handbook of financial time series*, pages 421–437, 2009.
- [32] Solym Manou-Abi. Approximate solution for a class of stochastic differential equation driven by stable processes. *Preprint hal-04217398*, 2023.
- [33] Solym Mawaki Manou-Abi. *Théorèmes limites et ordres stochastiques relatifs aux lois et processus stables*. PhD thesis, Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier), 2015.
- [34] H Masuda. Simple estimators for non-linear markovian trend from sampled data: I. ergodic cases. *MHF Preprint Series*, 7, 2005.
- [35] Remigijus Mikulevičius and Fanhui Xu. On the rate of convergence of strong euler approximation for sdes driven by levy processes. *Stochastics*, 90(4):569–604, 2018.
- [36] Elizbar A Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.
- [37] John P Nolan. *Univariate stable distributions*. Springer, 2020.
- [38] Olivier Menoukeu Pamen and Dai Taguchi. Strong rate of convergence for the euler–maruyama approximation of sdes with hölder continuous drift coefficient. *Stochastic Processes and their Applications*, 127(8):2542–2559, 2017.
- [39] Étienne Pardoux. *Probabilistic Models of Population Evolution: Scaling Limits, Genealogies and Interactions*, volume 1. Springer, 2016.
- [40] BLS Prakasa Rao. Asymptotic theory for non-linear least squares estimator for diffusion processes. *Statistics: A Journal of Theoretical and Applied Statistics*, 14(2):195–209, 1983.
- [41] BLS Prakasa Rao. Nonparametric estimation of linear multiplier in stochastic differential equations driven by stable noise. *arXiv e-prints*, pages arXiv–2109, 2021.
- [42] Gennady Samorodnitsky, Murad S Taqqu, and RW Linde. Stable non-gaussian random processes: stochastic models with infinite variance. *Bulletin of the London Mathematical Society*, 28(134):554–555, 1996.
- [43] Simon J Sheather. Density estimation. *Statistical science*, pages 588–597, 2004.
- [44] Yasutaka Shimizu. M-estimation for discretely observed ergodic diffusion processes with infinitely many jumps. *Statistical Inference for Stochastic Processes*, 9:179–225, 2006.
- [45] Yasutaka Shimizu and Nakahiro Yoshida. Estimation of parameters for diffusion processes with jumps from discrete observations. *Statistical Inference for Stochastic Processes*, 9:227–277, 2006.
- [46] Jian Wang. On the exponential ergodicity of lévy-driven ornstein–uhlenbeck processes. *Journal of Applied Probability*, 49(4):990–1004, 2012.
- [47] Geoffrey S Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372, 1964.
- [48] Chao Wei. Estimation for the discretely observed cox–ingersoll–ross model driven by small symmetrical stable noises. *Symmetry*, 12(3):327, 2020.
- [49] Wei Biao Wu. Nonparametric estimation for stationary processes. *University of Chicago. Technical Report*, 536, 2003.
- [50] Xu Yang. Maximum likelihood type estimation for discretely observed cir model with small α -stable noises. *Statistics & Probability Letters*, 120:18–27, 2017.
- [51] Xuekang Zhang, Haoran Yi, and Huisheng Shu. Nonparametric estimation of the trend for stochastic differential equations driven by small α -stable noises. *Statistics & Probability Letters*, 151:8–16, 2019.
- [52] Zhenzhong Zhang, Xuekang Zhang, and Jinying Tong. Exponential ergodicity for population dynamics driven by α -stable processes. *Statistics & Probability Letters*, 125:149–159, 2017.

6. APPENDIX: FIGURES AND TABLES

TABLE 1. Performance of the estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with a standard symmetric 1.6-stable driven OU process

n and T	True parameters	Estimate Parameters	RMSE
500 and T=1	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.70, \hat{\beta}_n = 0.08, \hat{\rho}_n = 0.80$	1.55
500 and T=10	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.69, \hat{\beta}_n = 0.015, \hat{\rho}_n = 3.38$	8.34
1000 and T=1	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.55, \hat{\beta}_n = 0.11, \hat{\rho}_n = 1.2$	1.10
1000 and T=10	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.56, \hat{\beta}_n = 0.13, \hat{\rho}_n = 4.87$	8.43
2000 and T=1	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.6346, \hat{\beta}_n = 0.1146, \hat{\rho}_n = 0.9335$	0.77
2000 and T=10	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.6029, \hat{\beta}_n = 0.064, \hat{\rho}_n = 4.33$	6.91

TABLE 2. Estimated drift coefficients $(\hat{\theta}, \hat{\mu})$ with the exchange rates data

n	Estimate Parameters
1000	$\hat{\theta}_n = 1.08, \hat{\mu}_n = 0.93$
500	$\hat{\theta}_n = 1.06, \hat{\mu}_n = 0.95$
200	$\hat{\theta}_n = 1.028493, \hat{\mu}_n = 0.9674548$

TABLE 3. Estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with the exchange rates data using an unknown strictly standard stable driven OU process

n	Estimated drift coefficients
1000	$\hat{\alpha}_n = 1.98, \hat{\beta}_n = 1, \hat{\rho}_n = 13.21$
500	$\hat{\alpha}_n = 1.99, \hat{\beta}_n = 1, \hat{\rho}_n = 10.11$
200	$\hat{\alpha}_n = 1.99, \hat{\beta}_n = 1, \hat{\rho}_n = 6.17$

TABLE 4. Performance of the estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with a standard symmetric 1.6-stable driven CIR process

n and T	True parameters	Estimate Parameters	RMSE
500 and T=1	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.66, \hat{\beta}_n = 0.13, \hat{\rho}_n = 0.88$	1.17
500 and T=10	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.51, \hat{\beta}_n = 0.08, \hat{\rho}_n = 5.95$	8.24
1000 and T=1	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.63, \hat{\beta}_n = 0.217, \hat{\rho}_n = 0.98$	3.72
1000 and T=10	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.53, \hat{\beta}_n = 0.35, \hat{\rho}_n = 6.32$	15.36
2000 and T=1	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.61, \hat{\beta}_n = 0.07, \hat{\rho}_n = 1.03$	2.07
2000 and T=10	$\alpha = 1.6, \beta = 0, \rho = 1$	$\hat{\alpha}_n = 1.64, \hat{\beta}_n = 0.04, \hat{\rho}_n = 4$	11.13

TABLE 5. Estimated diffusion and scaling parameters $(\hat{\alpha}, \hat{\beta}, \hat{\rho})$ with the exchange rates data using an unknown strictly standard SCIR process with given parameter q .

n	q	Estimate Parameters
1000	q=2	$\hat{\alpha}_n = 1.80, \hat{\beta}_n = -0.33, \hat{\rho}_n = 67.89$
1000	q=1.99	$\hat{\alpha}_n = 1.78, \hat{\beta}_n = -0.31, \hat{\rho}_n = 70.57$
1000	q=1.95	$\hat{\alpha}_n = 1.78, \hat{\beta}_n = -0.33, \hat{\rho}_n = 71.59$
1000	q=1.9	$\hat{\alpha}_n = 1.77, \hat{\beta}_n = -0.35, \hat{\rho}_n = 75.83$
1000	q=1.8	$\hat{\alpha}_n = 1.72, \hat{\beta}_n = -0.37, \hat{\rho}_n = 87.36$
1000	q=1.6	$\hat{\alpha}_n = 1.59, \hat{\beta}_n = -0.28, \hat{\rho}_n = 130.70$
1000	q=1.5	$\hat{\alpha}_n = 1.53, \hat{\beta}_n = -0.20, \hat{\rho}_n = 163.02$
500	q=2	$\hat{\alpha}_n = 1.79, \hat{\beta}_n = -0.12, \hat{\rho}_n = 39.61$
500	q=1.99	$\hat{\alpha}_n = 1.78, \hat{\beta}_n = -0.12, \hat{\rho}_n = 39.99$
500	q=1.95	$\hat{\alpha}_n = 1.77, \hat{\beta}_n = -0.093, \hat{\rho}_n = 41.56$
500	q=1.9	$\hat{\alpha}_n = 1.75, \hat{\beta}_n = -0.06, \hat{\rho}_n = 43.74$
500	q=1.8	$\hat{\alpha}_n = 1.72, \hat{\beta}_n = -0.009, \hat{\rho}_n = 48.98$
500	q=1.6	$\hat{\alpha}_n = 1.59, \hat{\beta}_n = 0.08, \hat{\rho}_n = 70.13$
500	q=1.5	$\hat{\alpha}_n = 1.53, \hat{\beta}_n = 0.073, \hat{\rho}_n = 86.03$
200	q=2	$\hat{\alpha}_n = 1.69, \hat{\beta}_n = -0.31, \hat{\rho}_n = 25.56$
200	q=1.99	$\hat{\alpha}_n = 1.69, \hat{\beta}_n = -0.30, \hat{\rho}_n = 25.83$
200	q=1.95	$\hat{\alpha}_n = 1.67, \hat{\beta}_n = -0.25, \hat{\rho}_n = 26.98$
200	q=1.9	$\hat{\alpha}_n = 1.65, \hat{\beta}_n = -0.19, \hat{\rho}_n = 28.59$
200	q=1.8	$\hat{\alpha}_n = 1.60, \hat{\beta}_n = -0.008, \hat{\rho}_n = 32.42$
200	q=1.6	$\hat{\alpha}_n = 1.51, \hat{\beta}_n = 0.04, \hat{\rho}_n = 43.64$
200	q=1.5	$\hat{\alpha}_n = 1.44, \hat{\beta}_n = 0.08, \hat{\rho}_n = 53.86$

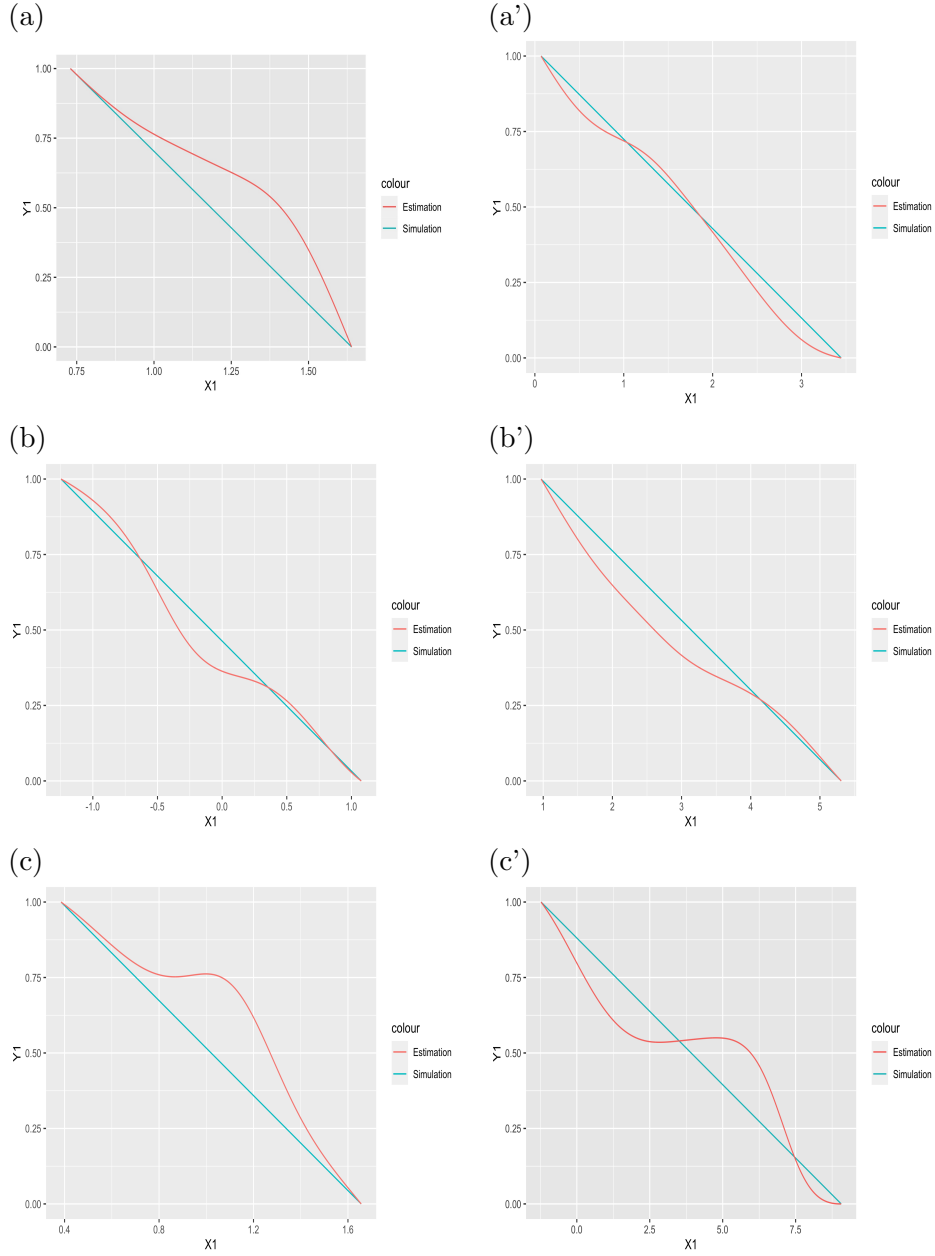
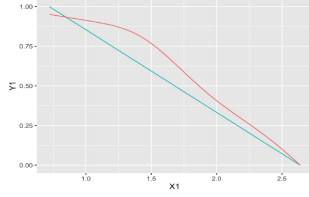
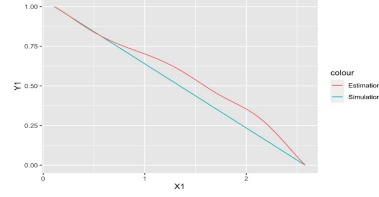


FIGURE 1. Performance of the Nadaraya-Watson kernel estimator with respect to the true drift in the case of a standard symmetric 1.6-stable driven OU process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

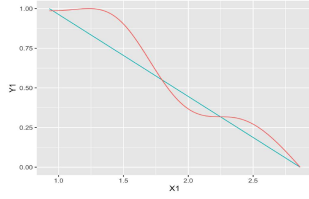
(a)



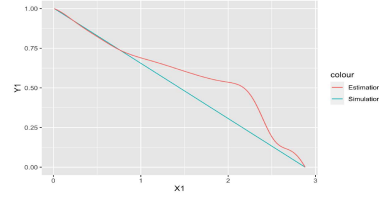
(a')



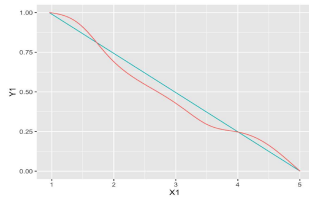
(b)



(b')



(c)



(c')

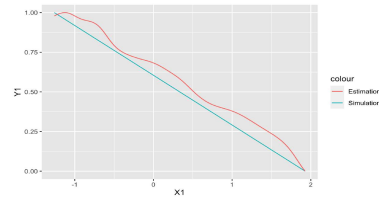
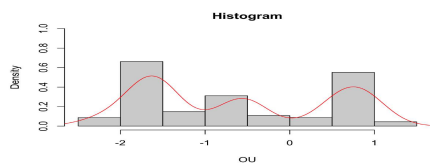
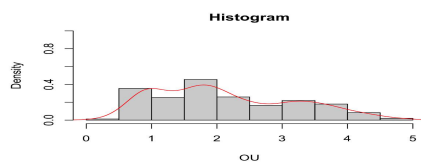


FIGURE 2. Performance of the Nadaraya–Watson kernel estimator with respect to the true drift in the case of a standard symmetric 1.6-stable driven CIR process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

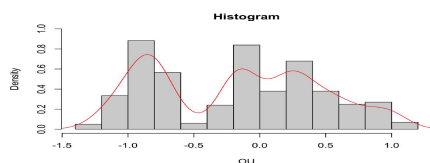
(a)



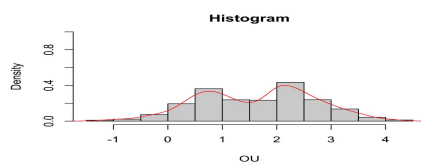
(a')



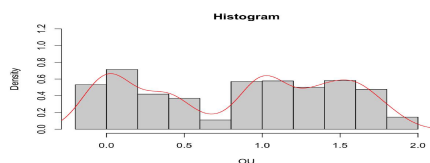
(b)



(b')



(c)



(c')

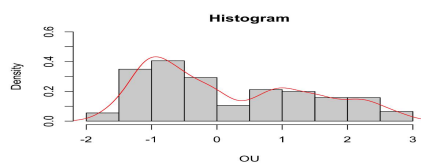
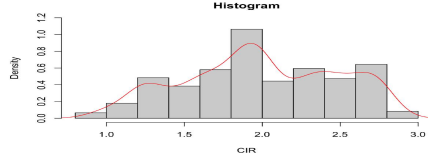
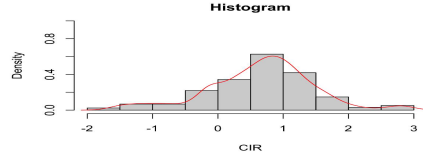


FIGURE 3. Kernel density estimation of a standard symmetric 1.6-stable driven OU process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

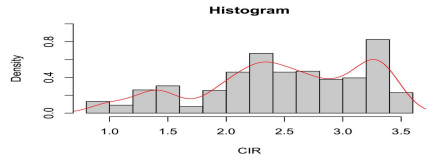
(a)



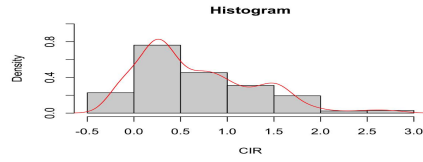
(a')



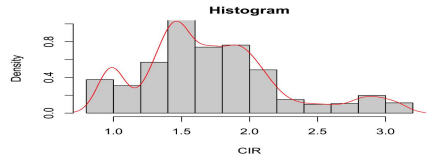
(b)



(b')



(c)



(c')

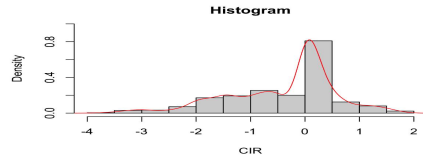
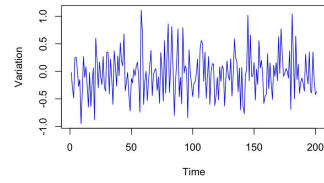
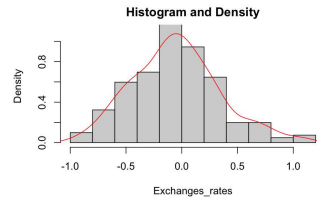


FIGURE 4. Kernel density estimation of a standard symmetric 1.6-stable driven CIR process with sample size $n = 500$ and $t = 1$ in (a), sample size $n = 500$ and $t = 10$ in (a'), sample size $n = 1000$ and $t = 1$ in (b), sample size $n = 1000$ and $t = 10$ in (b'), sample size $n = 2000$ and $t = 1$ in (c) and sample size $n = 2000$ and $t = 10$ in (c').

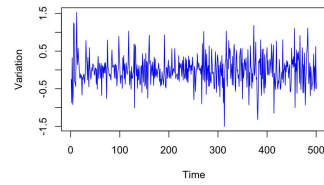
(a)



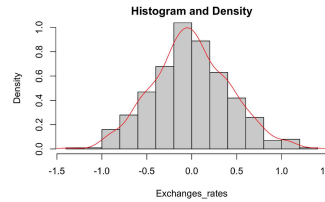
(a')



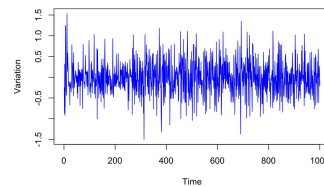
(b)



(b')



(c)



(c')

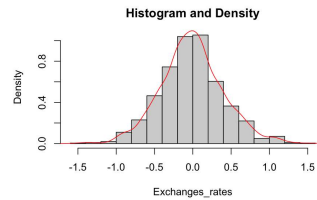
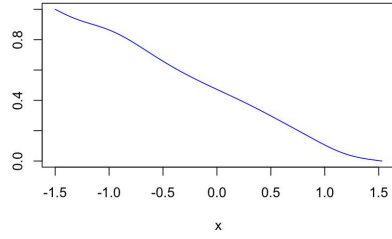
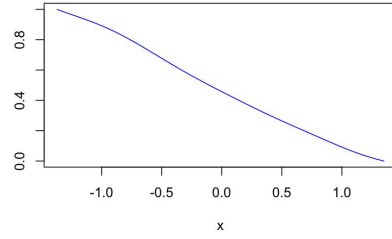


FIGURE 5. The exchange rates of the Canadian dollar against the US dollar global variation and density estimation with t sample size $n = 1000$ (a); local sample size $n = 500$ (b) and sample size $n = 200$ (c).

(a)



(b)



(c)

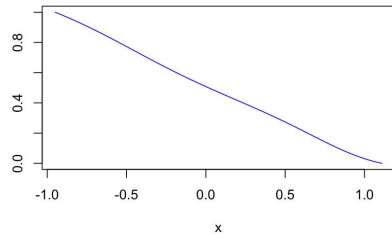
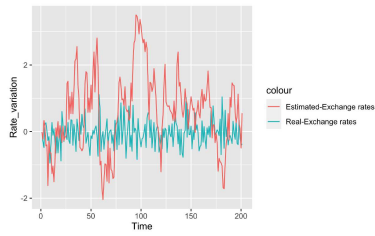
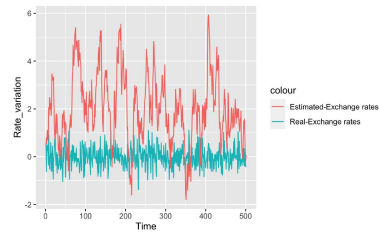


FIGURE 6. NW drift estimated function with the exchange rates of the Canadian dollar against the US dollar variation with the global sample size $n = 1000$ (a); local sample size $n = 500$ (b) and sample size $n = 200$ (c).

(a)



(b)



(c)

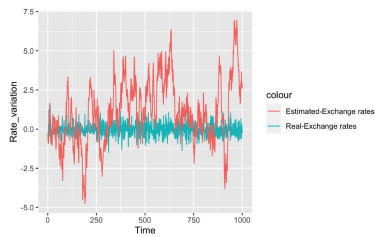


FIGURE 7. Comparison between exchange rates prediction with Stable driven OU process with sample size $n = 200$ (a); sample size $n = 500$ (b) and sample size $n = 1000$ (c).